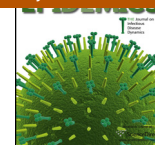




Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Virus-induced target cell activation reconciles set-point viral load heritability and within-host evolution

Anna Hool¹, Gabriel E. Leventhal¹, Sebastian Bonhoeffer*

Institute of Integrative Biology, ETH Zurich, Zurich, Switzerland

ARTICLE INFO

Article history:

Received 20 April 2013

Received in revised form 3 September 2013

Accepted 14 September 2013

Available online 23 September 2013

Keywords:

Set-point viral load

Virus-induced target cell activation

Within-host evolution

Between-host evolution

Viral load heritability

ABSTRACT

The asymptomatic phase of HIV-1 infections is characterised by a stable set-point viral load (SPVL) within patients. The SPVL is a strong predictor of disease progression and shows considerable variation of multiple orders of magnitude between patients. Recent studies have found that the SPVL in donor and recipient pairs is strongly correlated indicating that the virus genotype strongly influences viral load. Viral genetic factors that increase both viral load and the replicative capacity of the virus would result in rapid within-host evolution to higher viral loads. Reconciling a stable SPVL over time with high SPVL heritability requires viral genetic factors that strongly influence SPVL but only weakly influence the competitive ability of the virus within hosts. We propose a virus trait that affects the activation of target cells, and therefore viral load, but does not confer a competitive advantage to the virus. We incorporate this virus-induced target cell activation into within- and between-host models and determine its effect on the competitive ability of virus strains and on the variation in SPVL in the host population. On the within-host level, our results show that higher rates of virus-induced target cell activation increase the SPVL and confer no selective advantage to the virus. This leads to a build up of diversity in target cell activation rates in the virus population during within-host evolution. On the between-host level, higher rates of target cell activation and therefore higher SPVL affect the transmission potential of the virus. Random selection of a new founder strain from the diverse virus population within a donor results in a standing variation in SPVL in the host population. Therefore, virus-induced target cell activation can explain the heritability of SPVL, the absence of evolution to higher viral loads during infection and a large standing variation in SPVL between hosts.

© 2013 The Authors. Published by Elsevier B.V. Open access under [CC BY](http://creativecommons.org/licenses/by/4.0/) license.

Introduction

The course of an HIV-1 infection is divided into three stages characterised in part by their viral load. The viral load sharply increases during the primary infection period and then declines to reach a quasi-steady state level in the asymptomatic phase lasting a few years up to several decades before it increases again in the AIDS phase. During the asymptomatic phase the viral load fluctuates around a stable set-point viral load (SPVL) (Geskus et al., 2007). The magnitude of these fluctuations is small compared to the large variation of several orders of magnitude, which can be observed between patients (10^2 – 10^6 copies/ml) (Bonhoeffer et al., 2003; Fraser et al., 2007; Hockett et al., 1999; Mellors et al., 1996).

The SPVL is an important predictor of disease progression and a good proxy for virulence. Patients with higher SPVL progress faster towards AIDS (Mellors et al., 1996; Lavreys et al., 2006; Lyles et al., 2000) and have a higher chance per sexual contact to infect other people (Lingappa et al., 2010; Quinn et al., 2000). Therefore, understanding the mechanisms that shape SPVL is essential for the management of the disease.

Genome wide association studies have shown that host genetics and demography together can explain up to 22% of the variation in SPVL (Fellay et al., 2007, 2009). A number of recent studies have found a correlation in SPVL between donor and recipients, implying that also virus genetics strongly contribute to the variation in SPVL (Müller et al., 2011). The contribution of viral genetics to viral load variation is commonly quantified as the heritability of SPVL, which is defined as the proportion of variance in SPVL that is explained by variance in viral genetic factors (Visscher et al., 2008). Different methods in different patient cohorts quantified the heritability of SPVL and found estimates ranging from 0.2 to 0.6 (Müller et al., 2011; Alizon et al., 2010; Hollingsworth et al., 2010), although there is some inconsistency in the literature with regard to using correlation coefficient, the regression slope or the

* Corresponding author.

E-mail address: sebastian.bonhoeffer@env.ethz.ch (S. Bonhoeffer).¹ Equal contribution.

coefficient of determination as a measure of heritability (Müller et al., 2011). What virus genetic factors control viral load remains unclear. Higher viral load (i.e. counts of viral RNA copies) could arise from an increase in replicative capacity (Kouyos et al., 2011). In this case, virus genotypes that cause a higher viral load would have a competitive advantage over genotypes that cause a lower viral load.

The genetic diversity of HIV within a patient is large as HIV is prone to errors during replication (Overbaugh and Bangham, 2001; Rambaut et al., 2004). Furthermore, the virus population has a high turnover with a mean half-life of 1–2 days (Coffin, 1995; Ho et al., 1995; Perelson et al., 1996; Wei et al., 1995). The high heritability, the fitness differences between virus genotypes, the large virus population size within a host, the high mutation rate and the short generation time together lead to the expectation that within-host evolution should lead to higher viral loads over the long duration of the asymptomatic phase (Read and Taylor, 2001). Although SPVL increases slightly over the course of an infection, we do not observe a strong increase in viral load during the asymptomatic phase (Geskus et al., 2007).

One way to reconcile the contrasting observations of a high capacity for rapid evolution but the absence of strong within-host evolution towards higher viral load is to hypothesise that the genetic factors that control SPVL do not confer a competitive advantage to the virus on the within-host level. This is in contrast to viral traits that may influence SPVL (e.g. virion production, infectiousness, interactions with the immune system, or CTL escape mutations). These traits are beneficial and would thus be selected for during the course of an infection. The absence of large intra-host evolution, however, is an indication that these factors are unlikely the main drivers of the between-host diversity in SPVL.

One such factor that is selectively neutral but influence SPVL could be viral genes that contribute to the activation rate of target cell as activated target cells represent the pool of cells susceptible to infection. Combining analysis of clinical data and a very generic modelling approach Bonhoeffer et al. (2003) argued that activation rate of target cells may be a major contributor to variation in viral load between different patients. Clinical studies confirmed that higher SPVL is correlated with higher activation of target cells (Catalfamo et al., 2011) and that target cell activation is linked to faster target cell depletion, faster disease progression and higher transmission risk (Hazenberget al., 2003; Lawn et al., 2001). Biancotto et al. (2008) showed that activated target cells express activation markers such as CD25 or HLA-DR and are more susceptible to productive infection. While it is currently unclear how and to what extent the virus contributes to the rate of activation of target cells, there is no shortage of candidate factors (see Bartha et al., 2008). The population of activated target cells can be understood as a public good, i.e. all virus strains within the host benefit equally from this pool of susceptible cells regardless of how much the individual strains in the virus population contribute to the activation of target cells (Bartha et al., 2008; Brown, 1999). If all virus variants in the population benefit equally from the available pool of activated target cells, then the rate at which a virus strain activates target cells is selectively neutral (Bartha et al., 2008). Recently, Sanjuán et al. (Sanjuán et al., 2013) have shown that epitopes in HIV are more conserved when compared to HCV and argue that the activation of HIV-specific CD4 cells is therefore under positive selection. Here we refer to the activation of both HIV-specific and non-specific target cells as the majority of HIV infected cells are HIV non-specific (Douek et al., 2002).

If the virus-induced target cell activation is indeed a neutral (or nearly neutral) trait within the host, then we expect variation in the genetic factor influencing the activation rate to build up in a virus population within the host over the course of an infection given the rapid turnover of the viral population. On the

between-host level, however, virus-induced target cell activation influences viral fitness as a higher target cell activation rate would result in a higher SPVL. The virus population will evolve to a viral load that optimises the trade-off between the transmission probability and the duration of infection (Fraser et al., 2007; Alizon et al., 2009; Shirreff et al., 2011). We propose that this selection for maximal transmission results in the evolution of intermediate rates of virus-induced target cell activation. If the transmitted virus strain is randomly sampled from the virus population within the donor, then the within-host diversity in virus-induced target cell activation will lead to variation in SPVL around the optimum in the host population.

In this study, we incorporate virus-induced target cell activation into basic within-host and between-host models of HIV. We investigate the effect of virus strains with different activation rates on SPVL and their competitive ability within a host. We incorporate the findings from the within-host model into the between-host model to determine how the within-host diversity translates to variation in SPVL in the host population. In particular, we investigate whether models with virus-induced target cell activation are compatible with stable within-host SPVL, the observed variance in SPVL between hosts and the measured heritability.

Method

Within-host model

We extend the standard dynamical HIV model (Nowak and May, 2000) with an additional activation term of susceptible cells T that is proportional to the number of infected cells I ,

$$\frac{dT}{dt} = \lambda + \frac{2\lambda_I I}{1 + (I/K)} - \delta_T T - \beta T I, \quad (1)$$

$$\frac{dI}{dt} = \beta T I - \delta_I I. \quad (2)$$

Considering that the dynamics of free virus are much faster than those of the infected cells we assume that free virus and infected cells are in quasi-steady state (Bonhoeffer et al., 1997; De Boer and Perelson, 1998) such that the number of infected cells correlates with viral load. Susceptible cells are activated at a constant rate λ , are lost at rate δ_T and infected with rate βI . Infected cells are lost at rate $\delta_I > \delta_T$ to account for death due to infection. Susceptible cells are additionally activated at a rate $2\lambda_I I / (1 + I/K)$ that is proportional to the number of infected cells and the virus-determined parameter λ_I , and K is a constant. For $I \ll K$ the term approaches $2\lambda_I I$, for $I = K$ it is $\lambda_I I$ and for $I \gg K$ it saturates at $2\lambda_I K$.

To explore the competition in the within-host environment we developed a two-strain model distinguishing between virus strain 1 (I_1) with target cell activation $\lambda_{I,1}$ and virus strain 2 (I_2) with target cell activation $\lambda_{I,2}$,

$$\frac{dT}{dt} = \lambda + \frac{2\lambda_{I,1} I_1}{1 + (I_1 + I_2/K)} + \frac{2\lambda_{I,2} I_2}{1 + (I_1 + I_2/K)} - \delta_T T - \beta T I_1 - \beta T I_2, \quad (3)$$

$$\frac{dI_1}{dt} = \beta T I_1 - \delta_I I_1, \quad (4)$$

$$\frac{dI_2}{dt} = \beta T I_2 - \delta_I I_2. \quad (5)$$

Strains 1 and 2 differ only in the activation rate λ_I . Parameters and initial values are based on Althaus and De Boer (2011) and are rescaled to smaller population sizes for computational reasons (Table 1).

Table 1

Parameters and initial values for the within-host model. The values for λ , δ_T , δ_I and β are derived from Althaus and De Boer (Althaus and De Boer, 2011). β was calculated using the approximation $\beta(\text{per cell}) = \beta(\text{per virus})p/c$ with p being the burst size per infected cell and c the clearance rate of virus. t was lowered to account for additional virus-induced activation. λ and β are downscaled by the factor 1000 by which the population size was downscaled. T_0 is described by the uninfected equilibrium given by $\lambda i \delta_T$. I_0 is chosen big enough to avoid extinction due to stochastic events in the initial phase.

Parameter	Value	Explanation
T_0	1400	Initial value of uninfected target cells T [cells]
I_0	10	Initial value of infected target cells I [cells]
λ	140	Constant activation of target cells [cells d ⁻¹]
$\lambda_I, \lambda_{I,1}, \lambda_{I,2}$	0–0.9	Virus-induced activation of target cells [d ⁻¹]
δ_T	0.1	Net loss rate of uninfected target cells [d ⁻¹]
δ_I	1.0	Death rate of infected target cells [d ⁻¹]
β	0.003	Infection rate per infected target cell [d ⁻¹ cell ⁻¹]
K	1000	Carrying capacity [cells]

We developed both a deterministic and a stochastic version of the model. We solve the deterministic one-strain model for the equilibrium solution (T^∞ , I^∞) as a function of λ_I . The equilibrium solution of T^∞ is independent of λ_I ,

$$T^\infty = \frac{\lambda}{\delta_T} \quad \text{if } I_0 = 0, \quad (6)$$

$$\text{and } T^\infty = \frac{\delta_I}{\beta} \quad \text{if } I_0 > 0 \quad (7)$$

In contrast, I^∞ increases for increasing positive λ_I and asymptotically approaches zero as λ_I tends towards increasingly negative values,

$$I^\infty = \frac{1}{2\beta\delta_I} (-\delta_I\delta_T + \beta\lambda - \beta\delta_I K + 2\beta\lambda_I K + \sqrt{4\beta\delta_I(-\delta_I\delta_T K + \beta\lambda K) + (-\delta_I\delta_T + \beta\lambda - \beta\delta_I K + 2\beta\lambda_I K)^2}). \quad (8)$$

I^∞ represents the SPVL in our model.

In the stochastic two-strain model we further investigated the fate of a single mutant with $\lambda_{I,m}$ in a homogeneous resident population with $\lambda_{I,r}$. The analysis was conducted with three resident populations ($\lambda_{I,r} = 0.15, 0.45, 0.75$) and nine mutant types ($\lambda_{I,m} = 0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85$). We repeated the analysis for two different time points of introduction of the mutant: at the beginning of infection ($t_{in} = 0.01d$) and at the equilibrium stage ($t_{in} = 100d$). We repeated the simulation 15,000 times for $t_{in} = 0.01d$ and 20,000 times for $t_{in} = 100d$. The probability of fixation was calculated by dividing the number of simulations in which the mutant reached fixation by the total number of simulations where fixation of either type had happened before 1500d. More simulations are required when the mutant is introduced in the equilibrium state since the expected probability of fixation is lower.

Table 2

Parameters and initial values for the between-host model. Equations and parameters to calculate β and δ are taken from Fraser et al. (2007). I_0 is chosen big enough to avoid extinction due to stochastic events in the initial phase. Initial values for $\lambda_{I,pop}$ range between $R_0 \geq 0.75 R_{0,max}$.

Parameter	Value	Explanation
N	2000	Total population size (susceptible and infected hosts) [people]
I_0	90	Initial number of infected hosts [people]
$\lambda_{I,pop}$	–	Mean virus-induced target cell activation within an individual host
SPVL	Eqs. (6) and (7)	Set-point viral load of an individual host [copies ml ⁻¹]
β	Eq. (8)	Infectivity of an individual host [people ⁻¹ a ⁻¹]; described by $\beta_{max} = 0.317 \text{ y}^{-1}$, $\beta_{50} = 13,938 \text{ copies ml}^{-1}$ and $\beta_k = 1.02$
δ	Eq. (9)	Death rate of an individual host [a ⁻¹]; described by $D_{max} = 25.4 \text{ y}$, $D_{50} = 3058 \text{ copies ml}^{-1}$ and $D_k = 0.41$
σ_I	10–400	Within-host diversity in virus-induced target cell activation
σ_I'	0.1–0.7	

The models were implemented in R (R core Team, 2013). The deSolve package (Soetaert et al., 2010) was used for the deterministic model, the adaptivetau package (Johnson, 2012) for the stochastic model.

Between-host model

We model the between-host dynamics using a stochastic individual-based model in a host population of constant size $N = 2000$. Each host can either be susceptible or infected. The probability of being infected at a specific time point is the same for all susceptible hosts. Birth and death of susceptible hosts are not explicitly modelled. We assume that removed hosts are immediately replaced by new susceptible hosts immediately replace dying susceptible hosts. Parameter values and initial values of the variables are given in Table 2.

We only consider the asymptomatic phase of the infection such that the whole infection of a host can be characterised by its SPVL. The asymptomatic phase is expected to contribute most to the overall transmission potential due to its long duration (Hollingsworth et al., 2008). The relative contribution of the asymptomatic stage to the total transmission during HIV-1 infection was shown to be 71% in a serial monogamy scenario and 42% in a random mixing scenario, while the primary infection accounts for 9% and 31% respectively of the transmission in the two scenarios, and the late-stage infection for 20% and 27% respectively (Hollingsworth et al., 2008). The overall transmission potential across all stages is a function of SPVL (Fraser et al., 2007).

The mean virus-induced target cell activation of all the strains within a host, $\lambda_{I,pop}$, is the only specific characteristic assigned to an infected host. Because we do not observe significant changes in SPVL in patients we assume $\lambda_{I,pop}$ to be constant in a host over the whole duration of infection. Since it is unclear how genetic changes in the virus influence virus-induced target cell activation, we developed two different models that differ in the way $\lambda_{I,pop}$ translates into the SPVL of the host. We approximate Eq. (8) using a simplified function that translates $\lambda_{I,pop}$ into SPVL of a host,

$$\text{SPVL} = c(\lambda_{I,pop} + \sqrt{1 + (\lambda_{I,pop})^2}), \quad (9)$$

where we choose $c = 100$ resulting in realistic SPVL (I^∞ between 10^2 and 10^6 cells) for values of λ_I between 0 and 5000.

Following Fraser et al. (2007) the SPVL of a host determines both infectiousness of an individual host,

$$\beta = \frac{(\beta_{max}(\text{SPVL})^{\beta_k}) / ((\text{SPVL})^{\beta_k} + (\beta_{50})^{\beta_k})}{N}, \quad (10)$$

and the death rate,

$$\delta = \frac{(\text{SPVL})^{D_k} + (D_{50})^{D_k}}{D_{max}(D_{50})^{D_k}}. \quad (11)$$

Fraser et al. (2007) used patient data from two different cohorts to estimate best-fit curves for the relation between SPVL and

infectiousness, and SPVL and duration of the asymptomatic phase. The infectiousness is derived from data on the transmission rate within sero-discordant couples and was divided by the population size N for our model (frequency-dependent infection). The death rate δ is the inverse of the duration of the asymptomatic phase. The basic reproductive number R_0 is $\beta N/\delta$. The R_0 -curve is right-skewed and peaks ($R_{0,\max}$) at a SPVL of $10^{4.52}$ copies ml^{-1} (Fraser et al., 2007).

At transmission, the $\lambda_{I,\text{pop}}$ of the recipient is sampled from the population of viral strains in the donor. As it remains unclear how genetic variation translates into variation in SPVL we model the variation in $\lambda_{I,\text{pop}}$ within the donor using two different models. In the linear model we assume that mutational changes influence the SPVL in a linear fashion. The $\lambda_{I,\text{pop}}$ of the recipient is then drawn from a normal distribution with mean $\lambda_{I,\text{pop}}$ of the donor and standard deviation σ_I ,

$$\lambda_{I,\text{pop}}^{\text{recipient}} \sim N(\lambda_{I,\text{pop}}^{\text{donor}}, \sigma_I). \quad (12)$$

In the log-linear model the SPVL increases exponentially for increasing positive $\lambda_{I,\text{pop}}$,

$$\lambda_{I,\text{pop}}^{\text{recipient}} \sim 10^N (\log 10 \lambda_{I,\text{pop}}^{\text{donor}}, \sigma_I^*). \quad (13)$$

For large values of $\lambda_{I,\text{pop}}$, the log-linear model is similar to the next-generation matrix approach of (Shirreff et al., 2011).

We initialised the system by randomly choosing $\lambda_{I,\text{pop}}$ for I_0 infected individuals such that their $R_0 \geq 0.75 R_{0,\max}$. We conducted an analysis of the equilibrium in dependence of σ_I . The equilibrium state of the population is expected to be independent of the initial dynamics and we did not explore the initial dynamics in detail. For the linear model we conducted simulations with σ_I between 10 and 400 (10, 50, 100, 150, 200, 250, 300, 350, 400), for the log-linear model σ_I^* between 0.1 and 0.7 (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7). We simulated the population until it reaches equilibrium and recorded mean and variance of \log_{10} SPVL during the equilibrium phase of all infected individuals (longitudinal data), as well as the mean values of population mean and population variance of \log_{10} SPVL every 0.05 time steps (cross-sectional data). We also recorded the mean number of infected individuals in the equilibrium. The correlation coefficient of the \log_{10} SPVL between all successful donor-recipient pairs in the equilibrium is an estimate of the heritability of \log_{10} SPVL, h^2 . We repeated the simulation a 100 times for each σ_I and took the mean of the described measures over all replicates where the epidemic did not go extinct.

The model was implemented in R (R core Team, 2013) using the τ -leap method (Gillespie, 2001).

Results

Within-host dynamics

In the two-strain model, differences in β and δ_I lead to competitive exclusion of the strain with lower β or higher δ_I . In contrast, differences in λ_I have no effect on the fitness of individual strains and strains with different λ_I coexist. $^\infty$ in the two-strain model is equivalent to a one-strain model with $\lambda_I = (\lambda_{I,1} + \lambda_{I,2})/2$.

Single linear regressions for the fixation probability of the different mutants in a resident population with $\lambda_{I,r}$ and for the introduction time point $t_{\text{in}} = 0.01\text{d}$ showed for $\lambda_{I,r} = 0.15$ a intercept of 0.10 (95% CI [0.10,0.10], $p < 0.001$), for $\lambda_{I,r} = 0.45$ a intercept of 0.10 (95% CI [0.09,0.10], $p < 0.001$) and for $\lambda_{I,r} = 0.75$ a intercept of 0.10 (95% CI [0.09,0.10], $p < 0.001$). For the introduction time point $t_{\text{in}} = 100$, the single linear regressions showed for $\lambda_{I,r} = 0.15$ a intercept of 0.007 (95% CI [0.005,0.008], $p < 0.001$), for $\lambda_{I,r} = 0.45$ a intercept of 0.003 (95% CI [0.003–0.004], p -value < 0.001) and for $\lambda_{I,r} = 0.75$ a intercept of 0.001 (95% CI [0.001,0.002], $p < 0.001$). The

95% confidence intervals for the slopes included 0 in all six regressions.

We found that the fixation probability of a single mutant ($\lambda_{I,m}$) in a homogeneous resident population ($\lambda_{I,r}$) is well approximated by $1/N$ with N number of infected cells at the time point of introduction of the mutant. The fixation probability is independent of subsequent changes in population size, which is in agreement with the findings of Lambert (2006). If the mutant is introduced in the initial phase of the infection ($t_{\text{in}} = 0.01\text{d}$) the fixation probability is $1/I_0$ with I_0 the initial population size of infected cells (Fig. 1a). If introduced at the equilibrium stage ($t_{\text{in}} = 100\text{d}$) the fixation probability is $1/I^\infty$ with I^∞ the equilibrium population size of the resident population calculated from the deterministic model (Eq. (13)) with $\lambda_I = \lambda_{I,r}$ (Fig. 1b). The probability of fixation of a single new mutant is then smaller in resident populations with higher $\lambda_{I,r}$ (i.e. larger population size). The probability for fixation does not depend on the value of $\lambda_{I,m}$.

Between-host dynamics

We looked at the influence of the within-host diversity σ_I on the equilibrium state of the population. In both models, we observed extinctions of the epidemic for large σ_I (8/100 simulations in the linear model for $\sigma_I = 400$, 7/100 simulations in the log-linear model for $\sigma_I^* = 0.7$). The heritability h^2 of SPVL decreases with increasing σ_I from 0.98 ± 0.00 for $\sigma_I = 10$ to 0.65 ± 0.00 for $\sigma_I = 400$ in the linear model and from 0.95 ± 0.00 for $\sigma_I^* = 0.1$ to 0.72 ± 0.00 for $\sigma_I^* = 0.7$ in the log-linear model.

The longitudinal and cross-sectional mean and variance of \log_{10} SPVL in the infected population is shown in Fig. 2. The mean \log_{10} SPVL deviates more strongly from the optimal \log_{10} SPVL ($R_{0,\max}$) for lower heritability h^2 (Fig. 2a). The mean \log_{10} SPVL is lower in the cross-sectional data than in the longitudinal data since individuals with low SPVL have a longer duration of infection and are thus more likely to appear in a cross-section of infected individuals. Individuals with high SPVL have a higher turnover rate in the population due to their higher infectiousness and shorter duration of infection. Consequently, there are more individuals with high viral load when considering all infected individuals throughout the epidemic. Thus the longitudinal mean \log_{10} SPVL is both higher than the optimal \log_{10} SPVL and higher than the cross-sectional mean. The mean \log_{10} SPVL is generally higher in the linear model than in the log-linear model since choosing a new $\lambda_{I,\text{pop}}$ on a linear scale results in more variation in \log_{10} SPVL below the \log_{10} SPVL of the donor than above. Thus, given a donor with optimal \log_{10} SPVL, recipients that receive a \log_{10} SPVL below the optimum have a lower expected fitness than recipients that receive a \log_{10} SPVL above the optimum.

The variance in \log_{10} SPVL increases with decreasing heritability h^2 for both the linear and log-linear models and in both the cross-sectional and longitudinal data (Fig. 2b). In the linear model, the variance in the cross-sectional data is lower than in the longitudinal data since cross-sectional data favours low \log_{10} SPVL individuals, thus increasing the effect of choosing a new $\lambda_{I,\text{pop}}$ on a linear scale. In the log-linear model, we observed only a small difference in variance between the cross-sectional and longitudinal data.

The number of infected individuals decreases with decreasing h^2 from 701 ± 2 for $\sigma_I = 10$ to 132 ± 6 for $\sigma_I = 400$ in the linear model and from 643 ± 2 for $\sigma_I^* = 0.1$ to 131 ± 7 for $\sigma_I^* = 0.7$ in the log-linear model. The decreasing number of infected individuals implies a decreasing mean R_0 of the infection in the population due to a mean \log_{10} SPVL above (or below) the optimum \log_{10} SPVL and a increasing variance in the population, i.e. more people infected

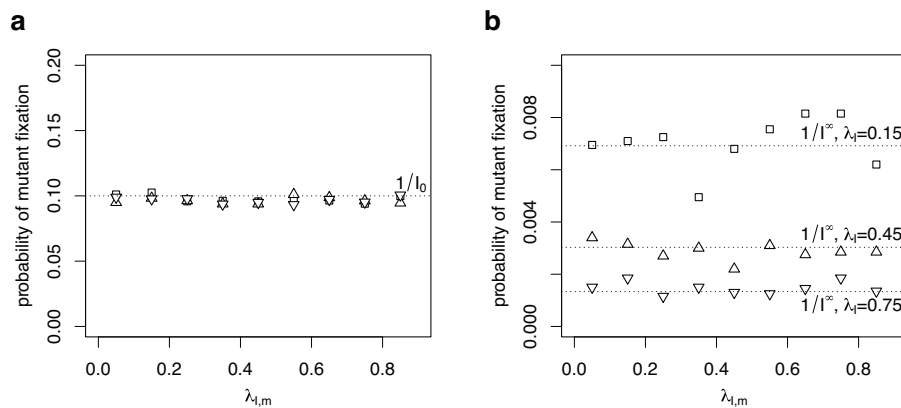


Fig. 1. Fixation probability of single mutants introduced at the beginning and in equilibrium. Single mutants with $\lambda_{i,r} = 0.15$ (\square), 0.45 (\triangle) and 0.75 (∇) are introduced into a homogeneous resident population. The probability of fixation of the mutant is well approximated by $1/N$ (dotted lines). (a) The mutant is introduced in the beginning of the simulation ($t_{i0} = 0.01d$). $N = I_0$ (I_0 = initial population size of infected cells). (b) The mutant is introduced at the equilibrium stage ($t_{i0} = 100d$). $N = I^\infty$ (I^∞ = equilibrium population size defined by $\lambda_{i,r}$). The confidence intervals (95%) of separate linear regression models all contain the $1/N$ line.

with a low-fitness virus. Additionally, a low number of infected individuals increases the risk for the disease to go extinct.

Discussion

Our results confirm that higher virus-induced target cell activation enhances the growth of the total virus population within a host and leads to higher SPVL. General activation of target cells benefits all virus genotypes within a patient equally in agreement with (Bartha et al., 2008). Thus virus strains that cause higher target cell activation do not have any fitness advantage and strains with different activation rates can coexist. Given the short generation time and the high mutation rate of the virus, we expect a fast build-up of standing within-host diversity with respect to target cell activation. Only the balance between neutral mutation and random genetic drift which is dependent on a realistic estimate of the effective population size influences which strain is dominant and how many strains coexist at the moment in the host.

In contrast to the SPVL, the uninfected target cell population within a host is constant and not affected by different rates of virus-induced target cell activation. This is in accordance with the observed T cell homeostasis in infected patients during the asymptomatic phase (Margolick et al., 1995). Higher target cell activation combined with T cell homeostasis is expected to lead to higher cell turnover rates and faster depletion of the available target cell pool, which then results in faster disease progression. If the virus can influence the activation rate of target cells, it will have a direct

effect on disease progression and the duration of the asymptomatic phase.

On the between-host level, the virus population evolves to a rate of virus-induced target cell activation that optimises its transmission fitness. This is compatible with Fraser et al. (Fraser et al., 2007) who suggested that the virus evolves to optimise the trade-off between infectiousness and duration of infection. A standing variation of SPVL between infected hosts is maintained by drift at transmission. More drift at transmission leads to a higher within-host diversity in virus-induced target cell activation. Increase in within-host diversity decreases heritability. We show that even high heritability values result in a considerable variation in SPVL in the infected host population and that the range variance in SPVL is compatible with the observed variance in real populations (Fraser et al., 2007).

The extent of the drift at transmission was also shown to determine the mean fitness of the disease in the host population and consequently the number of infected individuals.

We consider two different models of how mutational effects translate into variation in viral load. The linear model that uses a linear function treats virus-induced target cell activation as an additive trait, while the log-linear model using an exponential function applies a multiplicative understanding. We assumed constant within-host diversity in virus-induced target cell activation. The validity of this assumption, as well as the build-up of the within-host diversity needs to be further investigated in real patients. Continuous divergence from the founder strain and an increase and

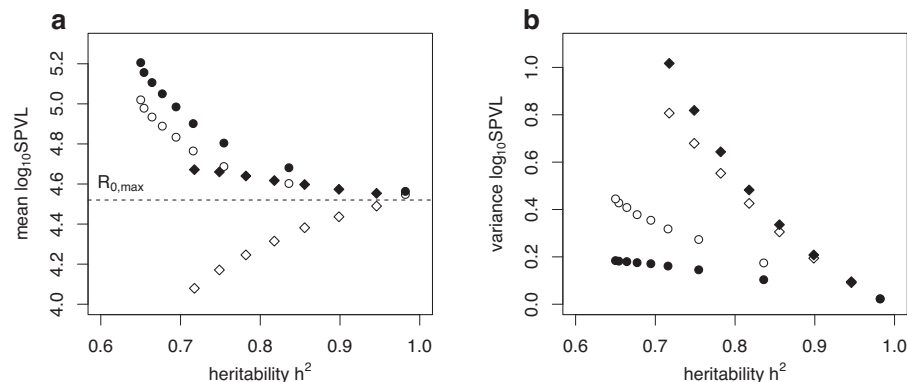


Fig. 2. Mean and variance in \log_{10} SPVL in the infected population against heritability. (a) Population mean \log_{10} SPVL; (b) population variance of \log_{10} SPVL. The different symbols are: linear model (circles), log-linear model (diamonds), longitudinal data (filled), cross-sectional data (empty). Heritability is measured as the correlation coefficient of \log_{10} SPVL of the donor–recipient pairs in the equilibrium stage for each $\sigma\lambda_i$. Error bars between replicas are smaller than the symbols and thus not shown.

saturation of diversity over time were observed in the HIV-1 *env* gene (Shankarappa et al., 1999) but changing selective forces during infection, as indicated by the changing ratio between synonymous and non-synonymous nucleotide substitutions in the HIV-1 *env* gene (Bonhoeffer et al., 1995), suggest that neutral diversity follows a different pattern than diversity in a trait undergoing selection.

We find heritability values that are larger than the heritability measured in previous studies (Müller et al., 2011). In our model, the only effect that decreases heritability is the drift event at transmission leading to differences in the mean virus genotype between donor and recipient. Host genotype or demography could also influence SPVL and therefore decrease heritability. In this case, we would assume that the same mean virus genotype is transmitted but that SPVL in the recipients differs from that of the donors. Day-to-day fluctuations in SPVL and measurement uncertainty can also decrease the heritability signal. A major question for further research is whether a trait which influences SPVL and is under weak selection on the within-host level (e.g. local target cell activation (Bartha et al., 2008) or replicative capacity) would show a similar pattern on the between-host level as we could show for selectively neutral general target cell activation. In this case, we would expect some standing within-host diversity (e.g. mutation-selection balance) but a changing mean during the infection due to selection. Such a trait would remain constrained at the between-host level due to differences in the transmission potential and would show a significant level of heritability.

Different mechanisms of how the virus can influence the activation of target cells have been proposed (Bartha et al., 2008). Identifying which viral factors are involved in virus-induced activation of target cells is important for a better understanding of interaction between SPVL and the pathogenesis of HIV. Genome wide association studies that map polymorphisms in the viral genome onto markers of immune activation are a promising first step towards identifying such viral factors.

Virus-induced target cell activation has an effect on the set-point viral load, is heritable between infections, is selectively neutral on the within-host level and is fitness-relevant on the between-host level. Given these characteristics, virus-induced target cell activation shows a way to reconcile high heritability, absence of within-host evolution and variation in set-point viral load in the infected population.

Acknowledgments

This research was supported in part by the European Research Council under the 7th Framework Programme of the European Commission (PBDR: Grant Agreement Number 268540) and in part by the Swiss National Science Foundation (Grant: 133129).

References

- Alizon, S., Hurford, A., Mideo, N., Van Baalen, M., 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* 22, 245–259.
- Alizon, S., von Wyl, V., Stadler, T., Kouyos, R.D., Yerly, S., Hirschel, B., Böni, J., Shah, C., Klimkait, T., Furrer, H., Rauch, A., Vernazza, P.L., Bernasconi, E., Battegay, M., Bürgisser, P., Telenti, A., Günthard, H.F., Bonhoeffer, S., the Swiss HIV Cohort Study, 2010. Phylogenetic approach reveals that virus genotype largely determines HIV set-point viral load. *PLoS Pathog.* 6, e1001123.
- Althaus, C.L., De Boer, R.J., 2011. Implications of CTL-mediated killing of HIV-infected cells during the non-productive stage of infection. *PLoS One* 6, e16468.
- Bartha, I., Simon, P., Müller, V., 2008. Has HIV evolved to induce immune pathogenesis? *Trends Immunol.* 29, 322–328.
- Biancotto, A., Iglehart, S.J., Vanpouille, C., Condack, C.E., Lisco, A., Ruecker, E., Hirsch, I., Margolis, L.B., Grivel, J.-C., 2008. HIV-1 induced activation of CD4+ T cells creates new targets for HIV-1 infection in human lymphoid tissue *ex vivo*. *Blood* 111, 699–704.
- Bonhoeffer, S., Holmes, E.C., Nowak, M.A., 1995. Causes of HIV diversity. *Nature* 376, 125.
- Bonhoeffer, S., Coffin, J.M., Nowak, M.A., 1997. Human immunodeficiency virus drug therapy and virus load. *J. Virol.* 71, 3275–3278.
- Bonhoeffer, S., Funk, G., Günthard, H., Fischer, M., Müller, V., 2003. Glancing behind virus load variation in HIV-1 infection. *Trends Microbiol.* 11, 499–504.
- Brown, S.P., 1999. Cooperation and conflict in host – manipulating parasites. *Proc. R. Soc. B* 266, 1899–1904.
- Catalfamo, M., Wilhelm, C., Tcheung, L., Proschian, M., Friesen, T., Park, J.-H., Adelsberger, J., Baseler, M., Maldarelli, F., Davey, R., Roby, G., Rehm, C., Lane, C., 2011. CD4 and CD8 T cell immune activation during chronic HIV infection: roles of homeostasis, HIV, type I IFN, and IL-7. *J. Immunol.* 186, 2106–2116.
- Coffin, J.M., 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* 267, 483–489.
- De Boer, R.J., Perelson, A.S., 1998. Target cell limited and immune control models of HIV infection: a comparison. *J. Theor. Biol.* 190, 201–214.
- Douek, D.C., Brenchley, J.M., Betts, M.R., Ambrozak, D.R., Hill, B.J., Okamoto, Y., Casazza, J.P., Kuruppu, J., Kunstman, K., Wolinsky, S., Grossman, Z., Dybul, M., Oxenius, A., Price, D.A., Connors, M., Koup, R.A., 2002. HIV preferentially infects HIV-specific CD4+ T cells. *Nature* 417, 95–98.
- Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Francioli, P., Mallal, S., Martinez-Picado, J., Miro, J.M., Obel, N., Smith, J.P., Wyniger, J., Descombes, P., Antonarakis, S.E., Letvin, N.L., McMichael, A.J., Haynes, B.F., Telenti, A., Goldstein, D.B., 2007. A whole-genome association study of major determinants for host control of HIV-1. *Science* 317, 944–947.
- Fellay, J., Ge, D., Shianna, K.V., Colombo, S., Ledergerber, B., Cirulli, E.T., Urban, T.J., Zhang, K., Gumbs, C.E., Smith, J.P., Castagna, A., Cozzi-Lepri, A., De Luca, A., Easterbrook, P., Günthard, H.F., Mallal, S., Mussini, C., Dalmau, J., Martinez-Picado, J., Miro, J.M., Obel, N., Wolinsky, S.M., Martinson, J.J., Detels, R., Margolick, J.B., Jacobson, L.P., Descombes, P., Antonarakis, S.E., Beckmann, J.S., O'Brien, S.J., Letvin, N.L., McMichael, A.J., Haynes, B.F., Carrington, M., Feng, S., Telenti, A., Goldstein, D.B., NIAID Center for HIV/AIDS Vaccine Immunology (CHAVI), 2009. Common genetic variation and the control of HIV-1 in humans. *PLoS Genet.* 5, e1000791.
- Fraser, C., Hollingsworth, T.D., Chapman, R., de Wolf, F., Hanage, W.P., 2007. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17441–17446.
- Geskus, R.B., Prins, M., Hubert, J.-B., Miedema, F., Berkhout, B., Rouzioux, C., Delfraissy, J.-F., Meyer, L., 2007. The HIV RNA setpoint theory revisited. *Retrovirology* 4, 65.
- Gillespie, D.T., 2001. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.* 115, 1716–1733.
- Hazenber, M.D., Otto, S.A., van Benthem, B.H., Roos, M.T., Coutinho, R.A., Lange, J.M., Hamann, D., Prins, M., Miedema, F., 2003. Persistent immune activation in HIV-1 infection is associated with progression to AIDS. *AIDS* 17.
- Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M., 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126.
- Hockett, R.D., Kilby, J.M., Derdeyn, C.A., Saag, M.S., Sillers, M., Squires, K., Chiz, S., Nowak, M.A., Shaw, G.M., Bucy, R.P., 1999. Constant mean viral copy number per infected cell in tissues regardless of high, low, or undetectable plasma HIV RNA. *J. Exp. Med.* 189, 1545–1554.
- Hollingsworth, T.D., Anderson, R.M., Fraser, C., 2008. HIV-1 transmission, by stage of infection. *J. Infect. Dis.* 198, 687–693.
- Hollingsworth, T.D., Laeyendecker, O., Shirreff, G., Donnelly, C.A., Serwadda, D., Wawer, M.J., Kiwanuka, N., Nalugoda, F., Collinson-Streng, A., Ssempija, V., Hanage, W.P., Quinn, T.C., Gray, R.H., Fraser, C., 2010. HIV-1 transmitting couples have similar viral load set-points in Rakai, Uganda. *PLoS Pathog.* 6, e1000876.
- Johnson, P., 2012. adaptivetau: Tau-leaping stochastic simulation.
- Kouyos, R.D., von Wyl, V., Hinkley, T., Petropoulos, C.J., Haddad, M., Whitcomb, J.M., Böni, J., Yerly, S., Cellerai, C., Klimkait, T., Günthard, H.F., Bonhoeffer, S., Swiss, H.I.V., Cohort Study, 2011. Assessing predicted HIV-1 replicative capacity in a clinical setting. *PLoS Pathog.* 7, e1002321.
- Lambert, A., 2006. Probability of fixation under weak selection: a branching process unifying approach. *Theor. Popul. Biol.* 69, 419–441.
- Lavreys, L., Baeten, J.M., Chohan, V., McClelland, R.S., Hassan, W.M., Richardson, B.A., Mandaliya, K., Ndinya-Achola, J.O., Overbaugh, J., 2006. Higher set point plasma viral load and more-severe acute HIV type 1 (HIV-1) illness predict mortality among high-risk HIV-1-infected African women. *Clin. Infect. Dis.* 42, 1333–1339.
- Lawn, S.D., Butera, S.T., Folks, T.M., 2001. Contribution of immune activation to the pathogenesis and transmission of human immunodeficiency virus type 1 infection. *Clin. Microbiol. Rev.* 14, 753–777, table of contents.
- Lingappa, J.R., Hughes, J.P., Wang, R.S., Baeten, J.M., Celum, C., Gray, G.E., Stevens, W.S., Donnell, D., Campbell, M.S., Farquhar, C., Essex, M., Mullins, J.I., Coombs, R.W., Rees, H., Corey, L., Wald, A., Partners in Prevention HSV/HIV Transmission Study Team, 2010. Estimating the impact of plasma HIV-1 RNA reductions on heterosexual HIV-1 transmission risk. *PLoS One* 5, e12598.
- Lyles, R.H., Muñoz, A., Yamashita, T.E., Bazmi, H., Detels, R., Rinaldo, C.R., Margolick, J.B., Phair, J.P., Mellors, J.W., 2000. Natural history of human immunodeficiency virus type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. Multicenter AIDS Cohort Study. *J. Infect. Dis.* 181, 872–880.
- Margolick, J.B., Muñoz, A., Donnenberg, A.D., Park, L.P., Galai, N., Giorgi, J.V., O'Gorman, M.R., Ferbas, J., 1995. Failure of T-cell homeostasis preceding AIDS in HIV-1 infection. The Multicenter AIDS Cohort Study. *Nat. Med.* 1, 674–680.
- Mellors, J.W., Rinaldo Jr., C.R., Gupta, P., White, R.M., Todd, J.A., Kingsley, L.A., 1996. Prognosis in HIV-1 infection predicted by the quantity of virus in plasma. *Science* 272, 1167–1170.

- Müller, V., Fraser, C., Herbeck, J.T., 2011. A strong case for viral genetic factors in HIV virulence. *Viruses* 3, 204–216.
- Nowak, M., May, R.M., 2000. *Virus Dynamics: Mathematical Principles of Immunology and Virology*. Oxford University Press, USA.
- Overbaugh, J., Bangham, C.R., 2001. Selection forces and constraints on retroviral sequence variation. *Science* 292, 1106–1109.
- Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., Ho, D.D., 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271, 1582–1586.
- Quinn, T.C., Wawer, M.J., Sewankambo, N., Serwadda, D., Li, C., Wabwire-Mangen, F., Meehan, M.O., Lutalo, T., Gray, R.H., 2000. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N. Eng. J. Med.* 342, 921–929.
- R Core Team, 2013. *R: A Language and Environment for Statistical Computing*.
- Rambaut, A., Posada, D., Crandall, K.A., Holmes, E.C., 2004. The causes and consequences of HIV evolution. *Nat. Rev. Genet.* 5, 52–61.
- Read, A.F., Taylor, L.H., 2001. The ecology of genetically diverse infections. *Science* 292, 1099–1102.
- Sanjuán, R., Nebot, M.R., Peris, J.B., Alcamí, J., 2013. Immune activation promotes evolutionary conservation of T-cell epitopes in HIV-1. *PLoS Biol.* 11, e1001523 EP.
- Shankarappa, R., Margolick, J.B., Gange, S.J., Rodrigo, A.G., Upchurch, D., Farzadegan, H., Gupta, P., Rinaldo, C.R., Learn, G.H., He, X., Huang, X.-L., Mullins, J.I., 1999. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol.* 73, 10489–10502.
- Shirreff, G., Pellis, L., Laeyendecker, O., Fraser, C., 2011. Transmission selects for HIV-1 strains of intermediate virulence: a modelling approach. *PLoS Comput. Biol.* 7, e1002185.
- Soetaert, K., Petzoldt, T., Setzer, R.W., 2010. Solving differential equations in R: Package deSolve. *J. Stat. Softw.* 33, 1–25.
- Visscher, P.M., Hill, W.G., Wray, N.R., 2008. Heritability in the genomics era – concepts and misconceptions. *Nat. Rev. Genet.* 9, 255–266.
- Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117–122.